It has been stated often in the past that practical acceptance of multi-valued logic awaits the development of suitable electronic means for fabricating a sufficiently powerful and general purpose logic. Though the search for such means has proceeded in parallel with every other development in the multi-valued field in the last thirty years, there is reason to believe that the pace is quickening and that success, by some measure, is sensibly upon us.

In support of this contention I offer in a general sense the all-pervasive presence of the integrated circuit industry in which formerly inconceivable complexity is a current low cost reality. Another, less obvious, property of this industry is its continuing search for complex, more powerful, functional entities to integrate. As linear dimensions of physically attainable devices increase, their accessibility (through traditional edge connections) rises in direct proportion. Obviously, however, the space for components rises only the square of the access function. This fact, coupled with a concurrent reduction in component scale through developments in photolithography, etc., greatly encourages the search for ever-increasingly-complex, less-access-limited products which use area in trade for perimeter. Such is one potential for multi-valued processing.

At the same time as these IC developments, and no doubt encouraged by them, there has been a rapid increase in the interest and accomplishment of circuit designers in the multi-valued area. A large variety of techniques and technologies have been exploited or are about to be. We have seen recent offerings, of direct interest to the multi-valued discipline, of the multi-emitter bipolar technology of Transistor-Transistor Logic (T2L), of variations of the technology of COS MOS and of developments relating directly to the current-switching designs of Emitter Coupled Logic (ECL). Shortly, we shall see evidence of the use of the recently developed Integrated Injection Logic (I2L) in multi-valued design.

It is with these perspectives that the present paper is written. It is important for the multi-valued discipline that each of the practitioners of its various facets are truly aware of, and interact with, the developments of the others. With this in mind I will begin an overview of electronic circuitry whose goal it is to enhance the appreciation of non-specialists in the rapid technological development which surrounds them.

We begin by reviewing some very basic properties of electronic devices, introduce a convention for drawing circuits which is a considerable aid in communicating and understanding, proceed to describe and interrelate global properties of available active (transistor) devices, then finally practice these basics on circuits typical of those you will see increasingly in the literature. The pace will be fast, the treatment unconventional, the scope distressingly broad. Our goal will be an improved appreciation of what has happened, and more importantly, an approach to understanding and appreciating more fully what is about to happen!

**Notation**

In the overview which follows, as well as in your subsequent attempts to fathom the mysteries of a circuit confronting you some time in the future, you will face a general problem of drawing conventions, notation etc., on which (unfortunately) there is some lack of universal agreement.

I will present, briefly, some facets of a notation, which I believe has growing acceptability and which, in any case, I believe can be used to redraw poorly presented circuits with a resulting improvement in understanding and communication.

The bases of the notation are several, namely:

1. Both signal and power supply aspects of a circuit should be represented since they are intimately coupled. That this is so may be appreciated better if it is realized that signals are, without exception, modulations or variations of power supply currents or voltages. This is really an energy conservation argument -- you don't get something for nothing; for every signal there must be power.

2. But, to the extent possible, signal and power functions should be kept as distinct as possible. This implies orthogonality of function will in fact be shown to correspond naturally to the two coordinate directions inherent in the drawing surface.

3. Draw everything by exception: if anything can be considered a global variable, define and use it. This is particularly important in digital circuits with characteristic repeated structures. If it really is the same as the rest, make it obvious -- learning time is too short to be squandered.

General results of deliberations motivated by the preceding thoughts are, to make a long story short, summarized below:

1. Consider the page to have a one-dimensional voltage gradient from top to bottom, positive at the top. The result will be that in components oriented vertically, current tends to flow from their top to their bottom. (Maybe you prefer a gravity analogy, or, heaven forbid, an anthropological one.) In any case, power supply voltages are thought to be supplied at the top and drip to the bottom. Building on this metaphor, the implied power supply (battery) is the percolator mechanism supplying gravitational potential for the drip process.

2. Signals will be represented in an orthogonal manner by flow across the page. While lines running vertically (on the page) are dominantly power supply oriented, lines running horizontally are signal oriented. Signal flow, as in writing (in some societies) is from left to right with a carriage return when the right edge is reached. There will of course be exceptions but this constitutes a reasonable basis for bringing order from chaos.

Specific symbols resulting from even more detailed deliberations of this kind are summarized in Figure 1. For completeness a reasonably full set of symbols are shown in their standard orientation. Many will remain mysterious for some of you for a while. Ideally this will be clarified eventually.

The power supply arrow and ground symbols are a very important mechanism for uncluttering the circuit diagram. Figure 2, using conventional battery symbols, shows the equivalent connection implied by the use of one upward directed arrow, one downward arrow and one ground or reference symbol. A more subtle important effect is that there is no longer any need for a horizontal line on the circuit whose sole function is to
Figure 1. Elements of a Drawing Convention for Electronic Circuits in Which Current Flows from Top to Bottom on the page. Note especially la, lb, and lg.

Figure 2. An Example of both the Power Supply Arrow Convention and the Uncluttered Diagram its Use Produces.

Figure 3. The Voltage-Current Relationship for a Linear Resistor.

Figure 4. A Simple Biasing Example with Resistors. The Current Generator Approximation.

Figure 5. The Role of the Capacitor in Establishing Time Scale.

Figure 6. Diodes -- Ideal and Practical Nonlinearities.
carry power. Conversely all lines across the page can be considered important as signal lines. This becomes very useful in a complex circuit and is a convention increasingly used.

The multiple arrow notation shown in Figure 1 is less essential, but handy. In this convention, more arrows represent a voltage of relatively greater magnitude. The "weight" attributed to each arrow used not be fixed and downward and upward directed symbols with equal numbers of barbs need not have the same magnitude.

**Linear Element Fundamentals**

**Passive Elements**

Resistance. A resistor is the fundamental power consuming element, used in a circuit often to control and limit currents from the battery to values suitable for modulation by active signal processing devices such as transistors. Ohm's law states that one volt impressed across one ohm produces a current of one ampere. More importantly, though not necessitated by this simple definition, many resistors are linear, and, for example, two volts across one ohm produces two amperes, etc. That is, in general, \( I = \frac{V}{R} \) where \( R \) is in ohms, \( V \) in volts, and \( I \) in amperes or some appropriately scaled units. In usual digital electronic circuits, \( V \) is measured in volts, \( R \) in kilohms or K and \( I \) is in milliamps or ma.

The general voltage/current relationship is shown plotted in Figure 3 for a 10 kilohm (10K) resistor for voltages from -20 to +20. Notice the perfect resistor is voltage polarity, or current direction, independent. At high voltages real resistors often depart from this relationship for many reasons.

Figure 4 shows a simple two-resistor circuit which allows us to introduce a concept which is very useful in the supply of power to active (transistor) elements to establish their region of operation, a process called biasing. In this circuit the 100 ohm (0.1K) resistor represents a transistor device, while the 10K ohm resistor is a biasing resistor intended to supply current \( I \). Their connection node, \( N \), is a place where signals pass as shown by the connector arrow symbolism. From the diagram:

\[ I = \frac{V}{R} = \frac{10V}{10K + 0.1K} = \frac{10V}{10.1K} = .99 \text{ ma or about 1 ma} \]

and the voltage

\[ V_N = IR = 0.99 \text{ ma (0.1K)} = 0.099 \text{ or about 0.1 volts} \]

Now in a usual (good) design the 10 volt power supply is relatively stable, varying perhaps as much as ±10% but often as little as ±1%. In either case, the total voltage appearing across the 10K resistor is relatively constant (within 10%) at about 10 volts, and the current \( I \) is equally constant at about 1 millamp. Under these conditions, and the assumption that the signals impressed at \( N \) do not violate the assumption (that the voltage across the 10K, \( V_{10K} \), is nearly always 10 volts), then the 10K can be replaced, at least conceptually, by a constant current generator of 1 milliamp value. This is shown also in Figure 4. The alternative symbol shown for the constant current generator is preferred (by me) since it reminds one that a real positive voltage supply is needed while showing, however, that its useful effect is to generate a relatively constant current.

**Capacitance.** A capacitor is the fundamental energy storage element of concern to the digital electronic designer. Its major attribute, that it is easy to fabricate (relative to inductance), is at one and the same time its most useful and least attractive feature. Everything one constructs brings capacitance, whether desired or not, and capacitance absorbs current and wastes time. It provides the dominant control of the time scale of operation of a logic circuit.

The voltage across a capacitor is the integral of the current flowing through it; that is \( v = \frac{1}{C} \int i \, dt \) with \( v \), the voltage across the capacitor in volts, \( i \), the capacitor current in amperes, \( C \), the capacitance in farads and \( t \), the time in seconds. Though the alternative form \( i = CV \) is perhaps more common and/or fundamental, the integral form is much more relevant to digital design. It shows that, as capacitance increases for a given current load, the voltage change required to activate a switching circuit requires a longer and longer time.

Either of the previous expressions, under suitable restrictions which happen to be relevant to digital design, can be rewritten as

\[ V = \frac{IT}{C} \]

or

\[ CV = IT \]

more usually

where \( V, I, T \) are respectively voltage, current, and time measured in some gross way.

Figure 5 illustrates this suitably for a logic application. Here \( C \) (10 picofarads, \( \times 10^{-11} \) farads or 10 pf) is a stray (wiring etc.) capacitance, \( I \) (1 ma) corresponds to the transistor device bias current, \( V \), (1 volt) is the voltage corresponding to the operating threshold of a subsequent logic stage and \( S \) is a switch corresponding to the operation of transistor device being examined. Since \( CV = IT \), the voltage \( V \) will reach its critical 1 volt threshold in time

\[ T = \frac{CV}{I} = \frac{10 \times 10^{-12} \times 1}{1 \times 10^{-3}} = 10 \times 10^{-9} \text{ seconds or 10 nanoseconds or 10 nsec.} \]

Notice, from a practical point of view, if we wished to reduce \( I \) to 0.1 ma to save power, the "operation time" would become 100 nsec. Alternatively if (separately) our subsequent logic was insensitive and needed 10 volts to operate, or if our wire was casual and his "rats nest" produced 100 pf, again \( T \) would be 100 nsec. In the dreadful event that all these effects appeared at once, the operation time would become 10 x 10 x 10 x 10 nsec. or 10 microseconds. Accordingly, beware!

**Nonlinear Elements**

The Semiconductor Junction Diode. Nonlinear elements are fundamental to establishing thresholds, and accordingly, to performing logic functions, while the fundamental nonlinearity is an ideal diode. Characteristics of two diodes, an ideal one and a "1 ma" silicon junction diode, are shown in Figure 6. Notice that the "positive" and "negative" scales are quite different. Though an ideal diode has, in the first quadrant, no "forward voltage drop", a real silicon junction has about 0.7 volts drop at a current which defines its capability. For example a so-called 1 ampere diode consists in fact of 1,000 1 ma diodes in parallel and has also a "forward voltage of about 0.7 volts. In the 3rd quadrant, where the voltage and current are reversed, the ideal diode current is zero while a real 1 ma diode might "leak" a current which is approximately constant at 1 nanoamp. Because of its construction, a 1 ampere diode would be expected to leak about 1 microamp.

Let us examine an important application of diodes to multivalued logic, namely generation of MAX and MIN functions. This is illustrated in Figures 7a and 7b for 2 input variables, \( A, B \) assumed for example to take on voltage values of 0, 1 and 2 volts. The one milliamp current source must exist to ensure that the diodes operate in the known part of their characteristic. It is a bias current, and might in an actual circuit consist simply of a resistor and a power supply voltage. To understand the operation, let us examine in
Figure 7. Diode Logic Gates - MAX and MIN Functions.

Figure 8. Three Terminal Active Devices - "Transistors"
detail only the MAX gate. Assume, at first, ideal diodes and that A is 2 volts while B is 0 volts. Now the constant current I tries to lower the voltage of C. If C were in fact connected to a capacitor, the voltage on the capacitor would fall slowly. But diodes D_A and D_B exist with characteristics shown in Figure 6. If, for some unspecified reason, the voltage at C were +10 volts, each ideal diode operates in quadrant 3 with reverse voltages of 8 and 10 volts respectively and zero current (since the diodes are ideal). Thus the voltage at C lowers (thinking of the capacitor ideal) until a voltage across the diodes decreases until, when the voltage at C reaches 2 volts, diode D_A reaches zero voltage and begins to conduct, entering quadrant 1. The total current of 1 ma begins to flow in it, forcing the voltage on C to stop at 2 volts. Thus an equilibrium is established with the voltage on A, V_A, or 2 volts, V_B = 0 volts and V_C = 2 volts. If now V_A falls to 1 volt, V_C will fall with it (slowly if a capacitor exists) until D_A begins to conduct again. If V_A subsequently goes to zero, so also will V_C, the current I being shared by both diodes. Correspondingly if V_B now goes to 2 volts say, it will force V_C to -2 volts. If a capacitor exists on the output, and the signal C is capable, large currents would flow through D_A, ultimately reducing to 1 ma established by I. Incidentally, the larger the current from B, the faster C will rise. Thus this circuit is operation-time-asymmetric; that is the output rise and fall times are distinctly different.

Summarizing, the net result is that C follows the most positive voltage on A or B. Thus for a "positive logic convention" this structure is a MAX gate. Under the same definition 7b shows a MIN gate. These roles of course reverse with a change to a negative logic convention.

But what happens with real diodes? By examining Figure 6, one can see that the effects are all approximately the same except that when a diode conducts, it loses 0.7 volts. Thus in Figure 7a when the most positive input is two volts the output will be 2.0 - 0.7 = 1.3 volts and the possible output values are accordingly 1.3, 0.3 and -0.7. Notice that the resulting output is incompatible, or at least subject to some question. If, for example, we had chosen our input logic values to correspond to +1.4, +0.7 and 0, the resulting outputs would be 0.7, 0.0 and -0.7. Thus, though partially compatible, the signals have also been partially down cycled.

The conclusion is that even a single diode logic level loses signal integrity and that of course two levels would be worse, particularly as the radix increases.

What is the solution? Three approaches suggest themselves. The first is to increase the signal separation so that the 0.7 volt diode drop "disappears" or, more correctly, is negligible. For example a choice of values of 14, 7 and 0 volts would make the uncertainty only 10%. Though this makes sense (sort of) in a radix 5 system, clearly troubles arise for a higher radix, say 10. More bad, is the effect on speed. Simply stated, for fixed currents (1 ma) and stray capacitance (10 pf), a 10 fold increase in voltage levels produces a 10 fold increase in speed. Thus this is not entirely a good solution.

What else? Two solutions exist to this problem. One is, by unspecified cunning means, to cancel the diode drop; the second is to provide some magic to reinterpret the output signals and regenerate them. Both solutions imply amplifiers. Amplifiers imply transistors. Transistors imply a new section.

Active Devices

Transistors. As implied by the collection of symbols used in Figure 1, there are many varieties of transistors. A large number of cryptic symbols confronts the casual observer: one sees reference to BJT, FET, IGFET, MOS, etc. etc. As well, since each of these comes in two flavors, P and N, describing the positive or negative polarity of internal carriers, a variety of variants are commonplace, including PNP or NPN and PMOS or NMOS or even CMOS, including both the previous and emphasizing their complementary properties.

Fortunately things are simpler than they appear to be. There are some unifying principles. It really is possible to learn enough of these to wade through this topic, as, hopefully we will see.

Three Terminals -- A Canonic Form. First, each of these devices has essentially only three connections -- a reference terminal, a control terminal, and an output terminal. Some of the more learned of you may object to the count 3 as gross oversimplification. Not so! All multiterminal semiconductor devices are really combinations of the basic three terminal one(s), with internal interconnections and additional leads brought out. The apparently new, exotic semiconductor device is really a combination of the basic 3 terminal ones to be described. Once a modest mastery of the basics is in hand (head?), a simple plea to the bounding narrowly trained specialists for the appropriate variable substitution, is all that is needed.

Three Terminal Notation. Let us return to the first principle enunciated, namely that all devices have three terminals, with each an implicit identified role, one as reference, one as control, and another as output. Figure 8 shows a block representation of such a device with corresponding devices of one conductivity type (N for negative carriers) placed alongside for reinforcement. Each of the spatially corresponding elements corresponds in function. Each has been labelled with the name evolved from its historical past.

It is interesting to note in the figure, where time proceeds to the right in general, the evolution of names of corresponding functions. One can see for example, in terms of the control function, the sequence grid, base and gate; in the output function are respectively control, gate, base, grid and reference, source, emitter, cathode and drain, collector, anode. In what follows I will use many of these interchangeably but will concentrate on the functional terms where available. Though the FET names will be used occasionally, a mixed trio, namely reference, gate and collector may often appear. This latter, apparently arbitrary choice comes from the fact that the control function of gating is clearly important for logic functions, that the gating threshold properties operate with respect to a reference, which also functions as a source of information (charge) carriers which are collected and used at the collector or output terminal of the device (See Figure 9, left part for a mixed equivalence).

1See Footnote at the end.
Figure 9. Three Terminal 'Transistor' Device Notation and Switching Oriented Models.

Figure 10. Direct Coupled Transistor Logic (DCTL) and its Switch Model.
A Transistor Device Model. In our search for basic principles governing transistor action, let us look at the device more closely in terms of the switching model of Figure 9. The rightmost two parts of the figure are a logic operation symbolism, using both logic gate and switch symbols to convey the basic idea of a transistor device represented, thus far, only by words. In this figure the triangular element is a logic gate, having, as logic must, some threshold above which the gate is on, below which it is off. The logical need for a reference to which "above" and "below" refer is represented by the wire leaving the lower part of the triangular logic gate and connected to the Reference terminal. This logic gate controls an electronic switch represented by the switch symbol (an open fence gate if you wish) in Figure 9c. Repeating, when the input signal exceeds a threshold, the gate(s) close(s), and connection is made from the source (of information carriers), represented by the Reference, to the Output (a collector of carriers).

The sketch in Figure 9d is simply an easier way to draw the switch used by relay logic designers, where the X represents the open shears used to cut the wire, leaving it open. Figure 9d is included to reinforce the equivalence to a differential amplifier input analyzer used for those who care about such concepts.

A Circuit Example -- DCTL. Let us practice our knowledge of transistors using this model for an ancient but honourable binary logic family called DCTL (direct coupled transistor logic). A simple but typical circuit is shown in Figure 10a using 2 BJTs which, for the present purposes, have a threshold voltage of 0.7 volts. In terms of the equivalent symbolism in Figure 10b this means that when Input B, hence G1, exceeds 0.7 volts more positive than R1, then the switch S1 closes connecting C1 to R1 and hence, by virtue of wiring, R2 to ground. If then (or already, or subsequently) A exceeds 0.7 volts, G2 exceeds R2 by 0.7, S2 closes and C2 is connected to R2 and hence, via S1 to ground. Thus the circuit output is connected to ground and, incidentally, 3 ma flows from the power supply through the 1K (load) resistor to ground. Now alternatively, if A is held at zero, S2 remains open and C stays at 3 volts, being connected there by the resistor. Alternatively if B is at ground, S1 is open and C cannot be grounded. Even though A is held at 0.7 volts or above, and accordingly G2 is potentially above R2 by the threshold voltage, R2 is not connected to anything which provides current. If in fact S2 closes, connecting R2 to 3 volts which is at 3 volts then R2 (S1 being open) would tend to rise to 3 volts. Since the assumption was 0.7 volts or so on C2 from A this tendency would turn off S2, leaving R2 as it was.

Incidentally this degenerative, negative feedback action of R2 with respect to G2 is an import property of all active transistor devices which accounts for the operation of "follower" circuits -- emitter follower, cathode follower, etc. of which you may have heard. In actual use, in the follower mode, the voltage at R2 would hover by an amount approximately equal to the threshold, just below the voltage on G2 while the cyclic hunt and seek process described above proceeds at the blinding electronic speeds. In the operation of R2 rises slightly, so must R2. If the voltage on G2 lowers (with minor circuit additions) so must R2. That is R2 follows G2. If, as is often true, G2 is used as an output, the signal on R2 is in phase with that on G2. In logic terms, there is no inversion produced.

Returning to our main stream, what does the circuit of Figure 10 do? Reviewing, we see that to lower the voltage on C (to ground) the voltages on both A and B must tend to attain voltages higher than the thresholds of each transistor. If either input, A or B, is at zero, the output, C, is high. The result, of course, is a logic operation which is either NAND or NOR, depending on the correspondence of logic "1" with positive or zero voltages, respectively.

The logic inversion represented by N of the compression of NOT AND to NAND is a very important property for logic design, as all logic designers know. More crucially, for some of the present purposes, it is also an inherent property of each and every transistor device you will see when used in the mode in which the gate, G, receives input signals and the collector, C, is used as output. As you recall from the previous development, if G serves as input and reference R is used as an output, results, having no inversion. As may be seen, there is time for example, the example of Figure 10 also demonstrates that signals input to R, while G is held fixed, reappears at C with no inversion. This is the mode taken on, part of the time, by T2 of Figure 10 relative to signals from T1.

Device Imperfections. Though the presentation, thus far, of transistor devices has been somewhat idealized (as some of you may note with disdain) it is adequate for understanding of a large number of transistor binary circuits, particularly those employing insulated gate field effect devices operating in the enhancement mode (more about which later) including a majority of MOSFET applications incorporating NMOS, PMOS and CMOS devices. The development thus far has reference furthermore to multi-valued logic and went to infinite valued logic, otherwise known as analog or linear electronics.

However transistors suffer a variety of imperfections not included thus far. Since the history of development of multi-valued logic is a story of incremental opportunism in which advantage is taken of the adversity of these device imperfections, it is important that we know what they are, or have been. The latter change of tense is important to note since, in fact, technology is forever striving to eliminate many of the imperfections on which ultimately unsuccessful multi-valued systems rest. On the positive side, however, each of these imperfections opens up wells of new material on which opportunism may thrive.

Output Imperfections. Since the imperfections of real transistors are many and varied, they will be presented in a somewhat minimal, constrained, list form to follow. The first list will concern itself with imperfections in the output mechanism of the transistor -- the switch in Figure 9.

1. Though the switch model implies equivalent bidirectional current flow between R (reference) and C (collector), this is precisely true only for most MOS devices, many FET devices and some BJT devices. Many devices have a preferred direction of current flow which can be accounted for by incorporating diodes (in series) in the switch model.

2. Following (1), there is perhaps the implication, and in reality there is the fact, that for some devices the role of R and C can be interchanged. This is true for a very large number of devices and results in a variety of alternative side effects of some use or misuse. Specifically emitters and collectors of BJT's are often exchanged for good or bad reasons, as are the source and drain of MOS devices. The possibility of exchange of source and drain in MOS is responsible for the existence of the important analog (and therefore most general) signal switch in the COS-MOS or CMOS technology.

3. The switch incorporated in each real transistor comes with two properties affecting current flow or, alternatively, affecting the voltage/current relations associated with the C and R terminals. Both properties depend furthermore on the voltage impressed between G and R. One of these is a simple series resistance accounting for the fact that as the current demanded by the load increases, the voltage across the switch increases.

4. When the load on the switch produces a voltage across it ranging from a few millivolts for a BJT
Figure 11. The Voltage-Current Relationship for a Typical MOS Transistor.

Figure 12. An Interpretation of Enhancement and Depletion Mode Operation.

Figure 13. Diode Transistor Logic (DTL).
to a few volts for an MOS device, a current limiting mechanism takes over. The result is that a current flows in the load which is constant, approximately independent of the voltage between C and R, but proportional in some sense to the voltage between G and R. Such an effect is shown in Figure 11 for an N channel, enhancement MOS transistor. The "constant" current and series resistance regions are delineated roughly. The (upward) slope of the "constant current" part is a second order imperfection associated with the transistor which shunts the gate when it conducts.

Further in some circuit applications the supply to the Reference terminal is not a voltage but a current source. In this case, since the Reference current is limited, so also is the current available at the collector, which becomes in turn a constant current generator.

**Imperfections at the Input.** Let us now turn our attention to detailed properties of the Gate input terminal of the transistor device characterized by Figure 9. Thus far we know of some "threshold voltage", measured with respect to the Reference, associated with opening and closing the output switch. When the Gate voltage is on one "side" of the threshold, the switch closes. When the voltage is on the "other" side, the switch opens.

A general unifying principle can be stated relative to which side is which: For all 3-terminal fundamental transistor devices, the direction in which the gate voltage must change to cause turn-on is that, with respect to the Reference, occupied by the voltage of the Collector terminal. An example of this can be seen with the BJT devices in Figure 10, where a +3 volt collector supply was used and a +0.7 volt threshold was assigned for turnon of the device. In this case, repeating for clarity, the threshold is 0.7 volts and an attempt to make the input voltage larger than this (say 0.8), in the direction of the 3 volt collector supply, turns the device on.

For a minimum device current gain of 30. Since the collector current of at least 0.7 ma, the base current is 1.3 - 0.7 = 0.6 ma, the current in R 1 is 5 - 1.4 = 1.3 ma as is the current in R 2 and a transistor inverter using T 1 and R 3. It is normally called a NAND gate in the DTL (Diode Transistor Logic) family.

Its operation may be understood briefly as follows. Provided A and B are both high, say 5 volts, node Y tries to rise toward 5 volts driven by current through R 1. As Y rises, diode D 3 tends to conduct with a 0.7 volt drop along its voltage, and the current in X, the current in R 1, is $I_{X} = I_{D3} = 1.3$ ma as is the current in D 1; the current in R 2 is $I_{R2} = 0.7$ ma, the base current is 1.3 - 0.7 = 0.6 ma permitting collector current of at least 30 (0.6) or 18 ma for a minimum device current gain of 30. Since the current required to lower C to zero is only 3 ma, the switch model of T 1 applies and the voltage at C is truly zero.

Considering the effect of lowering the voltage on (say) A, we see that when the voltage across D 1 reaches 0.7 directed along the arrow symbol, current flows, removing current from the base, lowering the voltage at X and turning off T 1. This effect occurs when the voltage at A goes below 0.7 + 0.7 - 0.7 or 0.7 volts. Thus the switching threshold of the entire NAND gate, at A or B, is 0.7 volts. Without Diode D 2 this threshold would be near zero and the logic element would not be self satisfying, having an input voltage range (below zero) beyond its output range (5 volts to near zero).

A third circuit example - T 2L

A related, more complex circuit, a T 2L (Transis-
Figure 14. Transistor Transistor Logic (TTL).

Figure 15. Emitter Coupled Logic (ECL).

Figure 16. Integrated Injection Logic (I^2L).

Figure 17. Complementary Symmetry Metal Oxide Semiconductor Logic (COSMOS or CMOS).
...tor Transistor Logic) NAND gate as shown in Figure 14 should be barely within our grasp at this moment. Its operation is similar to that of the circuit in Figure 13. Though apparently more complex, it offers a combination of ease of manufacture, improved speed and output power over that of the more basic DTL NAND. In Figure 14, transistor T2 is really two transistors with bases joined, collectors joined, and emitters separate.

When A and B are both up, at a voltage of 4 volts, say, the node Y tries to rise upward. With its emitters (A, B) more positive than its collector (X), NPN transistor T2 begins to operate backwards with X taking on reference or emitter terminal characteristics, notably that an internal diode, directed from Y to X, begins to conduct. Because T2 is operating backwards, its internal switch functions poorly and only a relatively small current flows from A or B to X. However all the current available from R1 flows through the diode in YX to X, raising the voltage at X. An equilibrium, consistent with what has preceded, is reached with T3 conducting, forcing C to zero, such that Z is at 0.7 with

\[ V_Z = 2.1 - 1.1 \text{mA} \text{ and the voltage at } W \text{ at } 0.7, \text{ the same as at } Z. \]

With Z at zero and W at 0.7 the total voltage shared by T4 and D4 is 0.7, inadequate to allow current through either. Thus T4 is cut off or not conducting. Thus if A and B are high, C is low.

When either of A or B descend by 0.7 below the voltage (2.1 volts) found at Y previously, transistor T2 reverts to its normal mode; it switches on, lowering the voltage at Y to 0.7 above the lower of A or B and holding the voltage at X equal to the voltage on the lower. Under these conditions, T1 and T2 cannot both conduct -- they switch off, allowing W to rise to 5 volts turning on T4 and D4. The net result is that the voltage on C is high at 5 - 0.7 - 0.7 or 3.6 volts when either of A or B is low (at zero). The switching threshold of the entire circuit is 1.4 volts in correspondence with the process outlined at the beginning of the paragraph.

A Fourth Example -- ECL

Another binary logic type illustrating a technique of considerable interest to multivalued logic designers is Emitter-Coupled Logic or ECL. Such a circuit implementing both OR and NOR functions is shown in Figure 15. Here signal levels center around the value -1.3 and are approximately -0.7 and -1.9, as will be seen. Provided A and B are both low (-1.9), neither T1 or T2 conduct and the current I flows entirely in T3. That is, this may be seen by computing the voltage on the emitter of T3 and rechecking its effect on the common connection to the emitters of T1 and T2. Specifically, the current I flowing in T3 will force a voltage of 0.7 across the emitter base pair (recall the follower example). Thus with its base at -1.3, the emitter of T3 will be at -2.0. Now considering T1 and T3, we see that each has only 0.1 volts between its base and emitter, well below the 0.7 volt threshold, and each must indeed be turned off, verifying our original estimation of the situation.

When current I flows from the emitter of T3 it automatically forces the switch within T3 to close and to extract this current from the collector. A current of 5 mA flowing in R3 = 240 ohms produces a voltage of -240 x 5 mA or -1.2 volts on F. Since T5 operates in the follower mode due to R5, the voltage on C follows to -1.2 - 0.7 or -1.9 volts.

When either one or both of inputs A or B rise to -0.7 volts, the associated transistor conducts, as a follower, with a voltage of -1.4 volts on F. The resulting voltage across T3 does not exceed the conduction threshold and T3 turns off, allowing F to rise to zero and C to rise to -0.7. The current I, meanwhile, flows from the collector of T1 or T2 to R3 to produce a voltage of -1.2 at 6 and -1.9 at D.

Thus if either A or B is high, C is high and D is also high. Thus C is an OR output while D is a NOR output.

A Fifth Example -- I2L

Another binary logic type illustrating a technique of considerable interest to multivalued logic designers is Injection Logic (I2L) as shown in Figure 16. It offers a variety of fabrication advantages including very small gate size and high packing density with very low, and controllable power consumption. In this circuit, T1, T2 and T3 are PNP BJT's in which the reference emitter terminal is at the positive, upper end of the device. As you will see all I2L devices have their bases grounded and emitters joined together to the (only) resistor R. This resistor forces current into the emitters, causing them to conduct with a voltage of +0.7 on each emitter. The resulting current \( \frac{I}{\text{R}} \) flows equally in each transistor since they are matched by virtue of the reference voltage being zero, while two volts is required for turnon. Thus C is forced to zero while T2 supplies current to charge the output capacitance. After C reaches zero, no current flows. This logic consumes no power while its inputs are static at normal logic levels.

If either A or B alone go to ground, the output stays at zero since one of T3 or T4 is off while one of T1 or T2 is on. If both A and B go to zero, T1 and T2 are turned off while both T3 and T4 are on, raising C to +5. Again, once C is high, no current flows. This circuit forms a positive logic NOR function.

A Sixth Example -- CMOS

A last binary example is shown in Figure 17, using Complementary Symmetry Metal Oxide Semiconductors, COSMOS, or CMOS for short. Here the MOS devices are drawn using the simplified MOS symbol. Each operates in the enhancement mode with a threshold of about 2 volts. Transistors T3 and T4 are P channel devices where the reference or source terminal is at the positive end of the device, in contrast with the N channel or negative reference transistors T1 and T2.

In this circuit, as we will see at the output, signal levels of +5 or 0 volts are sustained. When both inputs are at +5 volts, the thresholds of both T1 and T2 (+2 volts) are exceeded in the turnon direction while both T3 and T4 are cut off since their gate to source voltage is zero, while two volts is required for turnon. Thus C is forced to zero while T2 supplies current to charge the output capacitance. After C reaches zero, no current flows. This logic consumes no power while its inputs are static at normal logic levels.

Multivalued logic

At last! I am sure you feel it is about time, or has been for quite a while, even! A good example to begin with is the COSMOS circuit [2] of Figure 18 which intentionally bears a great resemblance to that of Figure 17. Only the power supplies have been changed to promote the innovation of replacing 2 logic values by 3 namely (+3, 0, -3).

For simplicity let us leave B at -3 volts and...
Figure 18. Ternary COSMOS.

Figure 19. Multivalued $I^2L$.

Figure 20. Ternary $T^2L$.

Figure 21. A Multivalued Storage Element Logic Representation.

Figure 22. A Five-Valued Storage Element Using ECL Techniques.
consider input A and output C only. Under these conditions, if B is off, then T2 is on, while T1 is off. When A is at ground, T2 is on, both T3 and T4 conduct while both T1 and T2 are off, hence X = C and Y = 0, as all ari are zero. The current (the interesting part) is A in the middle, at zero, the current thresholds of both T1 and T3 are exceeded; both are on. Now X rises to +3, Y lowers to -3, while C, driven by equal resistors, remains in the middle. The result, considering C and A alone, is that C is the ternary symmetrical complement of A, C = 1 - A, being opposite at the extremes and the same in the middle. Output X and Y form other useful inverses.

When the second input, B, is considered, the transistor connection can be seen to provide a MAX function, and in this case C becomes a ternary NAND of A and B.

A Second Multivalued Example -- High Radix 12L

The circuit of Figure 19 illustrates the possibilities inherent in reconbinations of binary [24]. For simplicity, and to avoid writing the first paper on the subject, I have shown only connections suitable for generating a four-valued current signal from a pair of two-valued current signals.

Here transistors T1, T2, and T3 are constant current generators characteristic of the I^2L technique. Transistors T4, T5 and T6 are the usual multicolonlector (or parallel interconnected) devices. The internal connection, shown in T4, for example, is novel, at least to this description, and is an example of a technique used commonly in linear circuits, called current mirroring. It is a negative feedback technique which forces an equality among currents. Specifically the current from T1 begins to flow to the base of T4 (provided A is open); the threshold of 0.7 volts is reached and T4 turns on. Because of the matched properties of the six transistors within T4, five (nearly) equal currents begin to flow in the collectors. Three of these currents are connected back to the base and are directed so as to steal the current supplied from T1, I, from the base of T4. In the equilibrium reached, the majority of the current is stolen and only a small residual flows to the base of T3. Thus by wiring, each collector is forced to conduct 1/3. Two of these collected, carry a current of 2/3 I when A is high, or zero, when A is low.

Similarly, the single wire from T5 to C carries 1/3 if B is high. At C three current sources are added: I from T2, 2/5 I or 0 from T4 and 1/3 or 0 from T6. The single mirror connection at T6 forces the sum, 1, 1/3, 2/3 I or 0, to appear at the output, depending on the signals A and B. The result is many things: It is at least a two bit D/A converter, but more importantly it is an example of great potential for high radix circuit implementation.

A Ternary NAND Using I^2L Components

Figure 20 shows a ternary NAND using commercial integrated circuit (IC) components in an unusual combination. [3] Transistors T1 and T2 are part of a gate expander IC while T3, T4 and T5 are part of an open-collector binary NAND package. Diode D1 and resistor R8 are discrete components. Resistor R2 and the 1.4 volt supply are not strictly necessary but are included here for tutorial purposes.

Consider the situation with input B positive or open. When input A is also positive (more than 1.4 volt, say) current from R1 flows out the collector of T1 (operating in an inverted fashion as did T5 in Figure 14) and this current of T1, T4 and T5 are all on. With P at 0.7 volts, forcing Q to ground. Meanwhile, since A is at 1.4 volts, resistor R8 raises N upward to turn on T4 (as a follower) which turns on T5. Node N reaches 0.7 volts while N attains 1.4. Resistor R4, incidentally, limits the current which flows through T1, from the +5 supply to T4. Transistor T5 being on, causes C to go to zero.

Now if the signal A is reduced to 0.7 volts, N is lowered to 0.7 since B is turned on, and both T4 and T5 turn off, allowing C to rise. Since A still now +3, however, so is P and T2 stays on, leaving Q at zero. Thus C rises only to 0.7 volts.

Now if the signal A is reduced to zero, P and N are reduced to zero and T2, as well as T4 and T5 turn off. Thus C rises to the supply or 1.4 volts. We see that with the ternary C = A but B is high. More generally when A and B are both active, C provides the NAND of A and B.

A Final Example -- A Current-Mode Five-Valued Storage Element

A final example has been chosen to illustrate many points, not least of which is that it may be possible at this time for you to begin to unravel large circuit schematics. The circuit to be discussed implements, in a radix-five, current-switching form, the storage element[6] logic diagram of Figure 21. Here the dot elements are MIN gates, plus elements are MAX gates while E is a level restoration or staircase element, consisting, logically of the cascade of 2 logic inverters. A clock C, with its complement E, samples the input variable X. The value of X is held in the positive feedback storage loop consisting of E, +, whose output is Y. A positive logic convention will be used in the implementation.

As you can see, the circuit of Figure 22, consists of a large number of NPN bipolar transistors, a few PNP BJTs and a few resistors, in direct proportion to their relative ease of manufacture. The arrow power supply notation has been used to eliminate a large number of explicit connections. A single upward-directed arrow is used to denote the (single) positive power supply connection while a single downward-directed arrow indicates the (single) relatively negative supply. Since, as we will see, the logic values are established relative to the upper supply, it is shown connected to ground, the signal reference. In addition, to further reduce apparent complexity, circled connector symbols, P, Q, R, S, denote wiring associated with connection to four voltage thresholds which serve, internally, to separate the values of a 5-valued logic system.

For completeness, all required components are shown explicitly. Thus, for example, though over half of the transistors (21 of 41) are used to generate constant currents, the temptation to use current generator symbols was turned aside. In the current generation circuits much use is made of the current-mirror linear-circuit technique described earlier in which many identical transistors are connected base to base and emitter to emitter, while current is forced to flow in one by connecting its collector to the common base to which a resistor and voltage supply is in turn connected. Since all transistors are identical, all connected with gate and reference terminals driven identically, the current supplied by each collector will be the same (as that forced through the resistor).

An explicit example of this includes T22 through T23 controlled by T24 and the resistor 6.5r. The apparent 4-leads connection to most of the transistors is a drafting artifact to make the parallel base connection more straightforward. Other currents are generated from T22 through T23 via T11 and 12r, from T24 via T7 and the same 12r, from T25 via T6 and a tapped 13r resistor, from T26 via T24 and 12r and finally from T27 through T23 and the same 12r.

But how large are these currents? Consider again T27 and the total number of current generation is on the order of 13r. During operation T27 operates with 0.7 volts from base to emitter and accordingly from collector to emitter as a result of the diode connection. With the 9.8 volt supply defined, a total voltage of 9.8 - 0.7 or 9.1 volts lies across the resistor string whose resistance totals 15r. Thus the current flow is 0.62 or 0.72
about 1 ma since r is 680 ohms. Also each unit r in the resistor string acquires $\frac{1}{2}$ or 0.7 volts drop. Since T32 through T43 are merely a string of followers all using the same 1 ma bias current derived through T46. Each of P, Q, R, S will lie 0.7 volts below the base voltage established by the resistor string. Thus the voltage on P will be 0.7 (4.5) - 0.7 or 3.85 volts below ground and Q, R, S will lie 0.7 volts progressively below this, as labelled in Figure 22.

Following similar arguments, one can establish other currents as shown in the figure, for example, 2ma or more precisely $\frac{3}{2}$ ma, in each of T29 through T32.

Let us return to the logic parts of the circuit of Figure 22. Transistors T7 and T6 with current sink T14 constitute an NPN follower MAX gate combining X and C to produce A. The connector name A has been included in Figure 21 for comparison. By definition, and in fact, A follows the most positive of X or C, with a 0.7 volt downward shift. Similarly B is formed via T8 and T10 from Y and E. Subsequently A and B are combined in a MIN gate, consisting of T27 and T28, to produce D. Notice that the 0.7 volt loss in the previous MAX gates is regained in the MIN and accordingly the signal levels at D are essentially the same at those at X and Y, the input and output terminals.

Likewise signals X and C are combined ultimately in the MIN gate consisting of T44 and T45 to produce E. Follower transistors T30 and T31 ensure a downward 0.7 volt shift to accommodate the net contribution of T44 and T45, ensuring that the logic values at E are compatible with those at Y as seen via the MAX gate using T9 and T19.

After processing by the follower input gates, the resulting signal D is compared against threshold voltages P, Q, R, S by differential pairs T34, T35 through T46, T47. These operate identically to similar components in binary ECL described earlier. For example, while D is a few tenths volts above P, transistor T34 is turned on while T35 acquires less than 0.7 from its base to emitter and stays off. The result is that the current extracted by T30 flows through T34 from a resistor r. At the same time since T35 is cut off, and, as we shall see, so are all of T37, T39 and T41, no current flows in the resistor r at W, and W reaches the upper power supply (0 volts). Accordingly the emitter of T16 is at 0.7, the base of T17 lower by (1.5x)$-\frac{1}{2}$ or 2.1 volts and Y, 0.7 volts lower again, at -0.7 - 2.1 = -3.8 volts. This is the most positive voltage level in the S valued system used, each of the others being 0.7 volts more negative than the last.

As the voltage at D goes below the threshold established at P of -5.85 volts, T35 and T43 exchange the 1 ma current drawn by T30 and W lowers by one signal interval or 0.7 volts, causing W to move from -3.5 to -4.2 volts. A further lowering of D causes switching of successive pairs and step reductions in W to -4.9, -5.6, and -6.3 volts.

While the Clock C is low, E is low, B follows Y, D follows E, (hence Y), and Y becomes a signal restored version of D. Thus the circuit 'remembers' any value established at Y by the input gates by virtue of its signal thresholding and regenerating properties.

With that we must terminate discussion, however brief, of a large circuit characteristic of the complexity and capability which modern IC technology brings to the multivalued logic arena.

In Conclusion

Let us finish this marathon with the wish that our rapid passage through the wilds of digital electronics leaves feelings more of hope than horror, hope that with a bit, or trit or whatever more practice, the mysteries of future multi-valued circuit developments will be revealed for all time.

Footnotes

1. As contrasted with the less desirable alternatives implied by the activities described by skin or, heaven forbid, founder.

2. Doubters more trained in the electronics than the average reader may be expected to be, may be mollified by cryptic reminders of common examples: The substrate connection of an MOS device is essentially the control terminal of a JFET series connected with the MOS device. An SCR, 4 layer diode, PUT, etc. is a simple connection of a PNP and an NPN BJT with resistors. Multi collector BJT's are simple base and emitter connected BJT's with separate collectors made available. Multi emitter BJT's are likewise parallel interconnected BJT's. A vacuum triode's heater leads merely supply a warm blanked to keep the organism alive in a cold world. Enough?

3. By relaxing the infinite gain idealization implied by Figure 10 and the accompanying description.

4. PNP BJT's.

5. Channel MOS, JFET.

6. The basic idea is due to Prof. P. Thompson, University of Ottawa, original inventor of T4L and is being implemented by N. Friedman and Professors P. Holmes and C.A.T. Salama at the University of Toronto.

7. r notation emphasizes the resistor ratio scaling properties of IC circuits. Ratios are relatively easy to establish while absolute values are not. Thus the r value of 680 ohms used here is not particularly critical.

References


Further Study

You may be interested in two small books which provide a basic, yet informative, treatment of this material. They are both authored by David Casasent, Associate Professor of Electrical Engineering, Carnegie-Mellon University, are entitled "Electronic Circuits" and "Digital Electronics" and are published by Quantum Publishers, Inc. (257 Park Avenue South, New York, N.Y. 10010). They each sell for US$5.95.